Evolving a Stigmergic Self-Organized Data Mining

Vitorino Ramos
CVRM-GeoSystems Centre
IST, Technical University of Lisbon, Portugal
vitorino.ramos@alfa.ist.utl.pt

Ajith Abraham
Department of Computer Science
Oklahoma State University, USA
ajith.abraham@ieee.org

Abstract – **Self-organizing complex systems typically are comprised of a large number of frequently similar components or events. Through their process, a pattern at the global-level of a system emerges solely from numerous interactions among the lower-level components of the system. Moreover, the rules specifying interactions among the system's components are executed using only local information, without reference to the global pattern, which, as in many real-world problems is not easily accessible or possible to be found. *Stigmergy*, a kind of indirect communication and learning by the environment found in social insects is a well known example of self-organization, providing not only vital clues in order to understand how the components can interact to produce a complex pattern but also can pinpoint simple biological non-linear rules and methods to achieve improved artificial intelligent adaptive categorization systems, critical for data mining applications. In this paper we attempt to show that a new type of data mining paradigm can be designed based on *Stigmergic* paradigms by combining the profit of several natural features of this phenomenon. We further propose hybrid bio-inspired swarm intelligence with evolutionary computation for an entirely distributed, adaptive, collective and cooperative self-organized data mining. The proposed framework is validated using Web mining data. Comparison of empirical results with other intelligent paradigms shows that the system presented is very promising.**

I. INTRODUCTION

Self-Organization refers to a broad range of pattern-formation processes in both physical and biological systems, such as sand grains assembling into rippled dunes, chemical reactants forming swirling spirals, cells making up highly structured tissues, and fish joining together in schools. As defined by *Camazine* et al. [29], self-organization is a process in which pattern at the global level of a system emerges solely from numerous interactions among the lower-level components of the system. Moreover, the rules specifying interactions among the system's components are executed using only local information, without reference to the global pattern.

One well-known example is provided by the emergence of self-organization in social insects, via direct mandibular, antennation, chemical or visual contact etc. or indirect interactions. The latter types are more subtle and defined by *Grassé* as stigmergy [11] to explain task coordination and regulation in the context of nest reconstruction in *Macrotermes* termites. An example [10], could be provided with two individuals, who interact indirectly when one of them modifies the environment and the other responds to the new environment at a later time. In other words, stigmergy could be defined as a typical case of environmental synergy or learning via the environment. *Grassé* showed that the coordination and regulation of building activities do not depend on the workers themselves but are mainly achieved by the nest structure: a stimulating configuration triggers the response of a termite worker, transforming the configuration into another configuration that may trigger in turn another (possibly different) action performed by the same termite or any other worker in the colony. Another illustration of how stigmergy and self-organization can be combined into more subtle adaptive behaviors is recruitment in social insects. Self-organized trail laid by individual ants is a way of modifying the environment to communicate with nest mates that follow such trails. It appears that task performance by some workers decreases the need for more task performance: for instance, nest cleaning by some workers reduces the need for nest cleaning [10,9]. Therefore, nest mates communicate to other nest mates by modifying the environment (cleaning the nest), and nest mates respond to the modified environment (by not engaging in nest cleaning); that is stigmergy. Division of labor is another paradigmatic phenomenon of stigmergy. But by far more crucial to the present work and aim, is how ants form piles of items such as dead bodies (corpses), larvae, or grains of sand (Figure 1). There again, stigmergy is at work: ants deposit items at initially random locations. When other ants perceive deposited items, they are stimulated to deposit items next to them, being this type of cemetery clustering organization and brood sorting a type of self-organization and adaptive behavior.

There are other types of examples (e.g. prey collectively transport), yet stigmergy is also present: ants change the perceived environment of other ants (their cognitive map, according to [6,28], and in every example, the environment serves as medium of communication. What all these examples have in common is that they show how stigmergy can easily be made operational. As mentioned by *Bonabeau* et al. [10], this is a promising first step to design groups of artificial agents which solve problems: replacing coordination (and possible some hierarchy) through direct communications by indirect interactions is appealing if one wishes to design simple agents and reduce communication among agents. Finally, stigmergy is often associated with flexibility: when the environment changes because of an external perturbation, the insects respond appropriately to that perturbation, as if it were a modification of the environment caused by the colony's activities. In other words, the colony can collectively respond to the perturbation with individuals exhibiting the same behavior. When it comes to artificial agents, this type of flexibility is priceless: it means that the agents can respond to a perturbation without being reprogrammed to deal with that particular instability. In a data mining context, this means that no classifier re-training is needed for any new data item types (new classes) arriving to the system, as is necessary in many classical models, or even in some recent ones. Moreover, the data-items that were used for supervised purposes in early stages of the colony evolution after further exploration of the search-space, can now, along with new items, be re-arranged through much better ways. Classification and/or data retrieval remains the same, but the system organizes itself in order to deal with new classes, or even new sub-classes. This task can be performed in real time, and in robust ways due to system's redundancy, as was shown in [23,2].

In this paper, we present a new type of data mining paradigm, which can be designed, based on stigmergic paradigms, by integrating the advantages of different natural features and incorporating them into a intelligent processing model. By hybridizing swarm intelligence with evolutionary computation, we seek for an entire distributed, adaptive, collective and cooperative self-organized data-mining model.

II. STIMERGIC WEB DATA MINING

Data clustering is also one of those problems in which real ants can suggest very interesting heuristics for computer scientists. One of the first studies using the metaphor of ant colonies related to the above clustering domain is due to *Deneubourg* [8], where a population of ant-like agents randomly moving onto a 2D grid are allowed to move basic objects so as to cluster them. This method was further generalized by *Lumer* and *Faieta* [18], (here after *LF* algorithm) applying it to exploratory data analysis, for the first time. In 1995, the two authors were then beyond the simple example, and applied their algorithm to interactive exploratory database analysis, where a human observer can probe the contents of each represented point (sample, image, item) and alter the characteristics of the clusters. They showed that their model provides a way of exploring complex information spaces, such as document or relational databases, because it allows information access based on exploration from various perspectives. However, this last work entitled "Exploratory Database Analysis via Self-Organization", according to [10], was never published due to commercial applications. They applied the algorithm to a database containing the "profiles" of 1650 bank customers. Attributes of the profiles included marital status, gender, residential status, age, a list of banking services used by the customer, etc. Given the variety of attributes, some of them qualitative and others quantitative, they had to define several dissimilarity measures for the different classes of attributes, and to combine them into a global dissimilarity measure (in, pp. 163, Chapter 4 [10]). More recently, *Ramos* et al. [26,25,28] presented a novel strategy (ACLUSTER) to tackle unsupervised clustering as well as data retrieval problems, avoiding not only short-term memory based strategies, as well as the use of several artificial ant types (using different speeds), present in those approaches proposed initially by *Lumer* [18]. Other works in this area include those from *Monmarché* et al. (1999 - [21]), *Ramos*, *Merelo* et al. [28,25,26,27], *Hoe* et al. [14], *Handl* and *Dorigo* [12] *Ramos* and *Abraham* [23,2,1]. Applications range from data clustering and exploratory data analysis to document processing and analysis [26,14,12], image retrieval systems [25], continuous mappings and classification [23], perception and memory on digital images [28], collective robotics [8,24], architecture and art [27], as well as WWW usage Mining [2] and network management [1]. Finally, and even if not related to clustering, other researchers as *Parpinelli* et al [22] approached Data-Mining via ACO, the famous *Dorigo's* Ant Colony Optimization Algorithm [10,9], extracting classification rules from data, having in mind an optimized framework.

Web mining is the extraction of interesting and useful knowledge and implicit information from artifacts or activity related to the WWW [17,7]. Web servers record and accumulate data about user interactions whenever requests for resources are received. User profiles could be built by combining users' navigation paths with other data features, such as page viewing time, hyperlink structure, and page content [13]. *Jespersen* et al [15] proposed a hybrid approach for analyzing the visitor click sequences. A combination of hypertext probabilistic grammar and click fact table approach is used to mine Web logs, which could be also used for general sequence mining tasks. *Mobaster* et al [20] proposed the Web personalization system, which consists of offline tasks related to the mining of usage data and online process of automatic Web page customization based on the knowledge, discovered. LOGSOM proposed by *Smith* et al [30], utilizes a self-organizing map (SOM) to organize web pages into a two-dimensional map based solely on the users' navigation behavior, rather than the content of the web pages. 'LumberJack' proposed by *Chi* et al [5] builds up user profiles by combining both user session clustering and traditional statistical traffic analysis using K-means algorithm. Also using SOM, *Merelo* et al [19] proposed the clustering and mapping of Web-log communities. *Joshi* et al [16] used relational online analytical processing approach for creating a Web log warehouse using access logs and mined logs (association rules and clusters). A comprehensive overview of Web usage mining research is found in [7,31].

In this paper, an ant colony clustering (ACLUSTER) [25] is proposed to segregate visitors and thereafter a linear genetic programming approach to analyze the visitor trends [2]. The results are compared with the earlier works using self organizing maps [32] and evolutionary-fuzzy C-means algorithm [3] to segregate the user access records and several soft computing paradigms to analyze the user access trends. Web access log data at the *Monash* University's Web site [33] were used for experiments. Average daily and hourly access patterns for 5 weeks (11 August'02 – 14 September'02) were used. In the subsequent section, the proposed self-organized and adaptive-evolving approach architecture is presented and experiment results are provided towards the end.

III. ANT CLUSTERING AND LINEAR GENETIC PROGRAMMING APPROACH (ANT-LGP)

The hybrid framework uses an ant colony optimization algorithm to cluster Web usage patterns. The raw data from the log files are cleaned and pre-processed and the ACLUSTER algorithm [25] is used to identify the usage patterns (data clusters). Each data-item is considered as an object, which will be transported by the artificial ants on a 2D classification space. The developed clusters of data are fed to a linear genetic programming model to analyze the usage trends. The swarm intelligence algorithm fully uses agents that stochastically move around the classification "habitat" following pheromone concentrations. That is, instead of trying to solve some disparities in the basic LF algorithm by adding different ant casts, short-term memories and behavioral switches, which are computationally intensive, representing simultaneously a potential and difficult complex parameter tuning, it was our intention to follow real ant-like behaviors as possible (some other features will be incorporated, as the use of different response thresholds to task-associated stimulus intensities, discussed later). In that sense, bio-inspired spatial transition probabilities are incorporated into the system, avoiding randomly moving agents, which tend the distributed algorithm to explore regions manifestly without interest (e.g., regions without any type of object clusters), being generally, this type of exploration, counterproductive and time consuming. Since this type of transition probabilities depends on the spatial distribution of pheromone across the environment, the behavior reproduced is also a stigmergic one. Moreover, the strategy not only allows guiding ants to find clusters of objects in an adaptive way (if, by any reason, one cluster disappears, pheromone tends to evaporate on that location), as the use of embodied short-term memories is avoided (since this transition probabilities tends also to increase pheromone in specific locations, where more objects are present). As we shall see, the distribution of the pheromone represents the memory of the recent history of the swarm, and in a sense it contains information, which the individual ants are unable to hold or transmit. There is no direct communication between the organisms but a type of indirect communication through the pheromonal field. In fact, ants are not allowed to have any memory and the individual's spatial knowledge is restricted to local information about the whole colony pheromone density. In order to design this behavior, one simple model was adopted by *Chialvo* and *Millonas* [6], and extended (as in [28]) due to specific constraints of the present proposal. As described in [6], the state of an individual ant can be expressed by its position r, and orientation θ. It is then sufficient to specify a transition probability from one place and orientation (r,θ) to the next (r^*,θ^*) an instant later. The response function can effectively be translated into a two-parameter transition rule between the cells by use of a pheromone weighting function (1). This equation measures the relative probabilities of moving to a cite r (in our context, to a grid location) with pheromone density $\sigma(r)$. The parameter β is associated with the osmotropotaxic sensitivity (a kind of instantaneous pheromonal gradient following), and on the other hand, $1/\delta$ is the sensory capacity, which describes the fact that each ant's ability to sense pheromone decreases somewhat at high concentrations. In addition to the former equation, there is a weighting factor $w(\Delta\theta)$, where $\Delta\theta$ is the change in direction at each time step, i.e. measures the magnitude of the difference in orientation. As an additional condition, each individual leaves a constant amount η of pheromone at the cell in which it is located at every time step t. This pheromone decays at each time step at a rate k. Then, the normalized transition probabilities on the lattice to go from cell k to cell i are given by P_{ik} [6] (2), where the notation j/k indicates the sum over all the pixels j which are in the local neighborhood of k. Δ_i measures the magnitude of the difference in orientation for the previous direction at time t-1.

$$W(\sigma) = \left(1 + \frac{\sigma}{1+\delta\sigma}\right)^{\beta} \qquad (1)$$

$$P_{ik} = \frac{W(\sigma_i)w(\Delta_i)}{\sum_{j/k} W(\sigma_j)w(\Delta_j)} \qquad (2)$$

In order to model the behavior of ants associated to different tasks, as dropping and picking up objects, we suggested [28] the use of combinations of different response thresholds. As we have seen before, there are two major factors that should influence any local action taken by the ant-like agent: the number of objects in his neighborhood, and their similarity (including the hypothetical object carried by one ant). *Lumer* and *Faieta* [18], use an average similarity, mixing distances between objects with their number, incorporating it simultaneously into a response threshold function. Instead, we recommend the use of combinations of two independent response threshold functions, each associated with a different environmental factor (or, stimuli intensity), that is, the number of objects in the area, and their similarity. Moreover, the computation of average similarities is avoided in the present algorithm, since this strategy can be somehow blind to the number of objects present in one specific neighborhood. In fact, in *Lumer* and *Faieta*'s work [18], there is an hypothetical chance of having the same average similarity value, respectively having one or, more objects present in that region. But, experimental evidences and observation in some types of ant colonies can provide us with a different answer. After *Wilson* (*The Insect Societies*, Cambridge Press, 1971), it is known that minors and majors in the polymorphic species of ants *Genus Pheidole*, have different response thresholds to task-associated stimulus intensities (i.e., division of labor). Recently, and inspired by this experimental evidence, *Bonabeau* et al. [10], proposed a family of response threshold functions in order to model this behavior. According to it, every individual has a response threshold θ for every task. Individuals engage in task performance when the level of the task-associated stimuli s, exceeds their thresholds. Author's defined s as the intensity of a stimulus associated with a particular task, i.e. s can be a number of encounters, a chemical concentration, or any quantitative cue sensed by individuals. One family of response functions $T_\theta(s)$ (the probability of performing the task as a function of stimulus intensity s, that satisfy this requirement is given by (Equation 3) [10], where n>1 determines the steepness of the threshold (normally n=2, but similar results can be obtained with other values of n>1). Now, at $s = \theta$, this probability is exactly ½. Therefore, individuals with a lower value of θ are likely to respond to a lower level of stimulus. In order to take account on the number of objects present in one neighborhood, (3) was used (where, n now stands for the number of objects present in one neighborhood, and $\theta = 5$), defining χ (4) as the response threshold associated to the number of items present in a 3 x 3 region around r (one specific grid location).

High-level description of ACLUSTER algorithm

/* **Initialization** */
For every object or data-item o_i **do**
Place o_i randomly on grid
End For
For all agents **do**
Place agent at randomly selected site
End For
/* **Main loop** */
For t = 1 to t_{max} **do**
For all agents **do**
$sum = 0$
Count the number of items n around site r
If ((agent unladen) and (site r occupied by item o_i)) **then**
For all sites around r with items present **do**
/* **According to Eqs. 4.4, 4.6 and Table 1 (4.1.1)** */
Compute d,χ, ε and P_p
Draw a random real number R between 0 and 1
If ($R \leq P_p$) **then**
$sum = sum + 1$
End If
End For
If (($sum \geq n/2$) or ($n = 0$)) **then**
Pick up item o_i
End If
Else If ((agent carrying item o_i) and (site r empty)) **then**
For all sites around r with items present **do**
/* **According to Eqs. 4.4, 4.5 and Table 1 (4.1.1)** */
Compute d,χ, δ and P_d
Draw random real number R between 0 and 1
If ($R \leq P_d$) **then**
$sum = sum + 1$
End If
End For
If ($sum \geq n/2$) **then**
Drop item o_i
End If
End If
/* **According to Eqs. 4.1 and 4.2 (section 4.1)** */
Compute $W(\sigma)$ and P_{ik}
Move to a selected r not occupied by other agent
Count the number of items n around that new site r
Increase pheromone at site r according to n, that is:
$P_r = P_r + [\eta + (n/\alpha)]$
End For
Evaporate pheromone by K, at all grid sites
End For
Print location of items
/* **Values of parameters used in experiments** */
$k_1 = 0.1$, $k_2 = 0.3$, $K = 0.015$, $\eta = 0.07$, $\alpha = 400$,
β=3.5, γ=0.2, $t_{max} = 10^6$ steps.

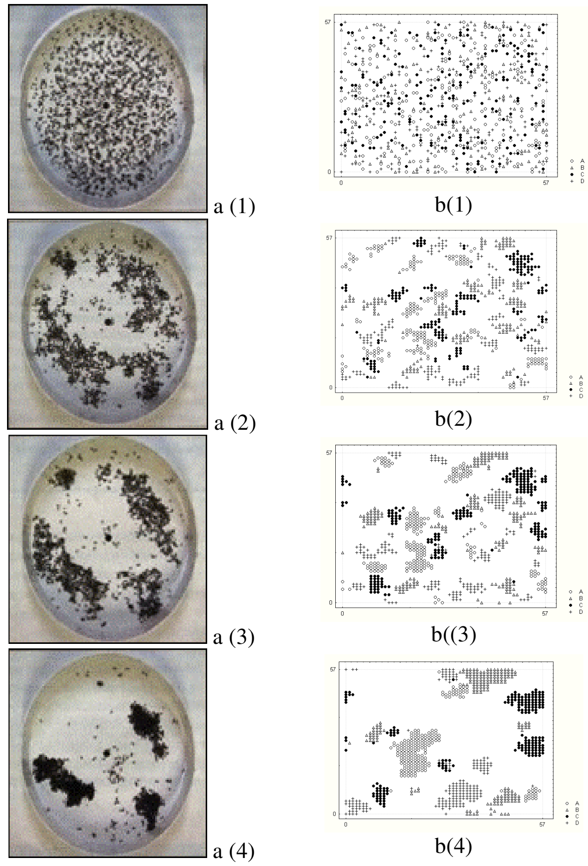

Fig. 1. – From a(1) to a(4), a sequential clustering task of corpses performed by a real ant colony. 1500 corpses are randomly located in a circular arena with radius = 25 cm, where *Messor Sancta* ant workers are present. The figure shows the initial state a(1), 2 hours a(2), 6 hours a(3) and 26 hours a(4) after the beginning of the experiment [10]. In b(1)-b(4) some experiments with the present algorithm, conducted on synthetic data (as in [25,18]). Spatial distribution of 800 items on a 57 x 57 non-parametric toroidal grid at several time steps. At t=1, four types of items are randomly allocated into the grid. As time evolves, several homogenous clusters emerge due to the ant colony action, and as expected the global entropy decreases [26].

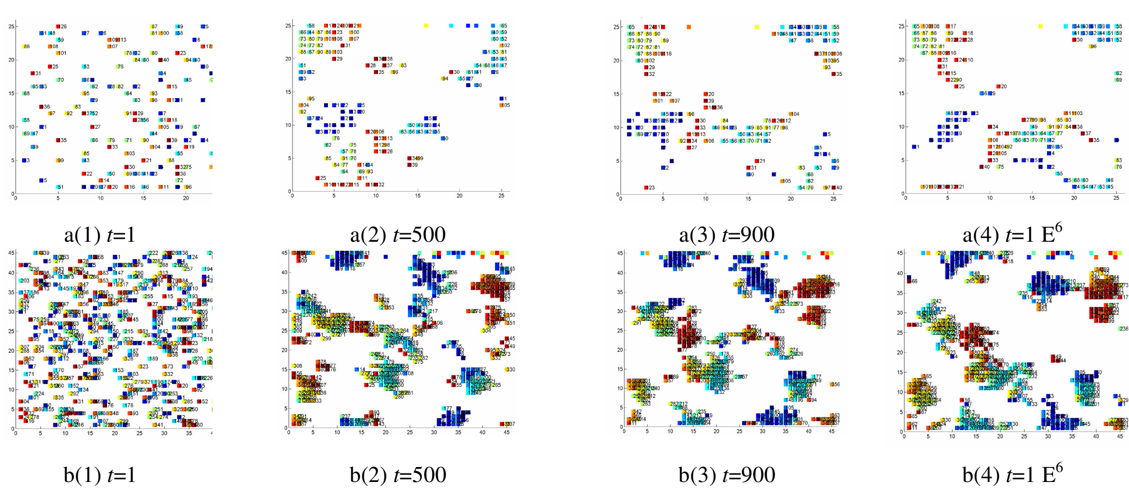

Fig. 2 – From a(1) to a(4) the snapshots represent the spatial distribution of daily Web traffic data on a 25 x 25 non-parametric toroidal grid at several time steps. At t=1, data items are randomly allocated into the grid. As time evolves, several homogenous clusters emerge due to the ant colony action. From b(1) to b(4), similar results for the hourly Web traffic data, on a 45 x 45 non-parametric toroidal grid with 48 ants present [2].

Now, in order to take account on the hypothetical similarity between objects, and in each ant action due to this factor, a Euclidean normalized distance d is computed within all the pairs of objects present in that 3 x 3 region around r. Being a and b, a pair of objects, and $f_a(i)$, $f_b(i)$ their respective feature vectors (being each object defined by F features), then $d = (1/d_{max}).[(1/F).\sum_{i=1,F}(f_a(i)-f_b(i))^2]^{½}$. Clearly, this distance d reaches its maximum (=1, since d is normalized by d_{max}) when two objects are maximally different, and d=0 when they are equally defined by the same F features. Moreover, δ and ε (5-6), are respectively defined as the response threshold functions associated to the similarity of objects, in case of dropping an object (5), and picking it up (6), at site r. Finally, in every action taken by an agent, and in order to deal, and represent different stimulus intensities (number of items and their similarity), present at each site in the environment visited by one ant, the strategy used a composition of the above defined response threshold functions (4-6). Several composed probabilities were analyzed [25] and used as test functions in one preliminary test. The best results were achieved with the test function #1 below (table 1), achieving a high classification rate (out of 4 different functions were used, as well the *LF* algorithm [18]. Alternatively, the system can also be robust feeding the data continuously (for instance, as they appear) as proven in past works [23]. For other algorithm details please refer [25,26,28].

$$T_\theta(s) = \frac{s^n}{s^n + \theta^n} \quad (3) \qquad \chi = \frac{n^2}{n^2 + 5^2} \quad (4)$$

$$\delta = \left(\frac{k_1}{k_1 + d}\right)^2 \quad (5) \qquad \varepsilon = \left(\frac{d}{k_2 + d}\right)^2 \quad (6)$$

TABLE 1. TYPE #1 PROBABILITY FUNCTIONS USED (AS IN [25,26])

Picking Probability	Dropping Probability
$P_p = (1-\chi).\varepsilon$	$P_d = \chi.\delta$

IV. EXPERIMENT SETUP AND RESULTS

After some preliminary analysis, we selected the statistical data comprising of domain byte requests, hourly page requests and daily page requests as focus of the cluster models for finding Web users' usage patterns. The most recently accessed data were indexed higher while the least recently accessed data were placed at the bottom. For each dataset (daily and hourly log data), the algorithm was run twice (for t=1 to 1E^6) in order to check if somehow the results were similar (which they appear to be, if we look into which data items are connected into what clusters). The classification space is always 2D, non-parametric and toroidal.

Linear genetic programming is a variant of the GP technique that acts on linear genomes [4]. Its main characteristics in comparison to tree-based GP lies in that the evolvable units are not the expressions of a functional programming language (like LISP), but the programs of an imperative language (like c/c ++). An alternate approach is to evolve a computer program at the machine code level, using lower level representations for the individuals. This can tremendously hasten the evolution process as, no matter how an individual is initially represented, finally it always has to be represented as a piece of machine code, as fitness evaluation requires physical execution of the individuals. Besides the inputs '*volume of requests*' and '*volume of pages (bytes)*' and '*index number'*, we also used the *'cluster information'* provided by the clustering algorithm as an additional input variable. The data was re-indexed based on the cluster information. Our task is to predict (a few time steps ahead) the Web traffic volume on a hourly and daily basis. We used the data from 17 February 2002 to 30 June 2002 for training and the data from 01 July 2002 to 06 July 2002 for testing and validation purposes. We used a LGP technique that manipulates and evolves a program at the machine code level. The settings of various linear genetic programming system parameters are of utmost importance for successful performance of the system. The population space has been subdivided into multiple subpopulation or demes. After a trial and error approach, the parameter settings used for the experiments were: *Population size:* 500, *Tournament size:* 4, *Maximum no. of tournaments:* 120.000, *Mutation frequency:* 90%, *Crossover frequency:* 80%, *Number of demes:* 10, *Maximum program size:* 512 and *Target subset size:* 100. These experiments were repeated three times and the test data was passed through the saved model and the empirical comparisons of the proposed framework with some of our previous works are depicted in Table 2 [1][3][32].

TABLE 2. PERFORMANCE OF PERIODIC WEB VISITOR PREDICTIONS

	Root Mean Squared Error (RMSE)			
	Daily		**Hourly**	
	Train	**Test**	**Train**	**Test**
ANT-LGP	0.0191	0.0291	0.2561	0.035
i-Miner	0.0044	0.0053	0.0012	0.0041
SOM-ANN	0.0345	0.0481	0.0546	0.0639
SOM-LGP	0.0543	0.0749	0.0654	0.0516

V. CONCLUSIONS

In this paper we presented a hybrid evolutionary and self-organized bio-inspired technique for data mining, unsupervised clustering and data exploratory analysis, while sketching a clear parallel between a mode of problem-solving in social insects and a distributed, reactive, algorithmic approach. Some of the mechanisms underlying corpse clustering, brood sorting and those that can explain the worker's behavioral flexibility, as regulation of labor and allocation of tasks have first been introduced. At the level of agent movement on the grid, a truly stigmergic model was introduced (section 4) in order to deal with clusters of objects, avoiding randomly moves which can be counterproductive in the distributed search performed by the swarm, and adopted by all past models. In fact, the present algorithm, along with [26,25], were the first to introduce pheromone traces on agents to determine random explorations and encourage objects cluster formation, a successful feature not implemented even in some recent proposals [14]. However, while achieving similar results compared to the *LF* model [18], as pointed by the spatial entropy of solutions at each iteration, the present algorithm is by far simpler. Moreover, for some of response thresholds compositions used, results are superior while using the present algorithm for the majority of time iterations [25], that is, entropy is always lower, even if at the end they tend to be the same value. As a final advantage, the present framework does not require any initial information about the future classification, such as an initial partition or an initial number of classes. Using ACLUSTER [23], a robust classifier could be achieved, which produces class decisions on a continuous stream data, allowing for continuous mappings. As we know, many categorization systems have the inability to perform classification and visualization in a continuous basis or to self-organize new data-items with respect to the older ones (even more, into new labels if necessary), unless a new training happens. This disadvantage is also present in self-organizing maps. The proposed ANT-LGP model seems to work very well for the problem considered. The empirical results also reveal the importance of using optimization techniques for mining useful information. Useful information could be discovered from the clustered data.

REFERENCES

[1] Abraham A., Business Intelligence from Web Usage Mining, Journal of Information & Knowledge Management (JIKM), 2(4), pp. 375-390, 2003.

[2] Abraham A. and Ramos V., Web Usage Mining using Artificial Ant Colony Clustering and Genetic Programming. In Proc. Congress on Evolutionary Computation, IEEE Press, pp. 1384-1391. 2003.

[3] Abraham A., i-Miner: A Web Usage Mining Framework Using Hierarchical Intelligent Systems. The IEEE Int. Conf. on Fuzzy Systems, FUZZ-IEEE'03, pp. 1129-1134, 2003.

[4] Banzhaf W., Nordin P., Keller E.R., Francone F.D., Genetic Programming: An Introduction on The Automatic Evolution of Computer Programs and its Applications, Morgan Kaufmann Publishers, Inc., 1998.

[5] Chi E.H. et al, LumberJack: Intelligent Discovery and Analysis of Web User Traffic Composition. Proc. of ACM-SIGKDD Workshop on Web Mining for Usage Patterns and User Profiles, ACM Press, Canada, 2002.

[6] Chialvo, D.R. and Millonas M.M., How Swarms Build Cognitive Maps. *The Biology of Intelligent Autonomous Agents*, 144, NATO ASI Series, pp. 439-450, 1995.

[7] Cooley R., Web Usage Mining: Discovery and Application of Interesting Patterns from Web Data. *Ph. D. Thesis,* Department of Computer Science, University of Minnesota, USA, 2000.

[8] Denebourg J.-L. et al, 1991. The Dynamic of Collective Sorting Robot-like Ants and Ant-like Robots. *Procs. of SAB´90 – 1st Conf. on Simulation of Adaptive Behavior: From Animal to Animats*, MA: MIT Press, pp. 356-365.

[9] Dorigo M. et al, 2000. Ant Algorithms and Stigmergy. Future Generation Computer Systems, 16, pp. 851-871.

[10] Eric Bonabeau, Marco Dorigo and Guy Théraulaz, 1999. *Swarm Intelligence: From Natural to Artificial Systems*, Santa Fe Institute in the Sciences of the Complexity, Oxford Univ. Press, New York.

[11] Grassé P.-P., 1984. *Termitologia,* Fondation des Sociétés. Construction. Tome II. Paris, Masson.

[12] Handl J. and Dorigo M., 2003. On the Performance of Ant-based Clustering. *Proc. of the 3rd Int. Conf. on Hybrid Intelligent Systems*, IOS Press, Dec. 03, Australia.

[13] Heer J. and Chi E.H., 2001. Identification of Web User Traffic Composition using Multi-Modal Clustering and Information Scent. Proc. Of the Workshop of Web Mining, SIAM Conf. on Data Mining, pp. 51-58.

[14] Hoe, K.M. et al, 2002. Homogeneous Ants for Web Document Similarity Modeling and Categorization. From Ant Colonies to Artificial Ants - Procs. of the 3rd Int. Work. on Ant Algorithms, Springer-Verlag, Sep. 02, Belgium.

[15] Jespersen S.E. et al, 2002. A Hybrid Approach to Web Usage Mining, Data Warehousing and Knowledge Discovery, LNCS 2454, Springer-Verlag, pp. 73-82.

[16] Joshi K.P. et al, 1999. Warehousing and Mining Web Logs. *Proc of the 2nd ACM-CIKM* Workshop on Web Information and Data Management, pp. 63-68.

[17] Kosala R. and Blockeel H., 2000. Web Mining Research: A Survey. *ACM SIGKDD Explorations*, 2(1), pp. 1-15.

[18] Lumer E.D. and Faieta B., 1994. Diversity and Adaptation in Populations of Clustering Ants. *Procs. of SAB´94 – 3rd Conf. on Simulation of Adaptive Behavior: From Animal to Animats*, The MIT Press/Bradford Books, Cambridge, MA.

[19] Merelo J.J. et al, 2004. Clustering Web-based Communities using Self-Organizing Maps, *Proc. of IADIS Conf. on Web Based Communities*, Portugal, 2004.

[20] Mobaster B. et al, 1999. Creating Adaptive Web Sites through Usage-based Clustering of URLs. *Proc. Of the Workshop on Knowledge and Data Engineering Exchange*, USA, pp. 19-25.

[21] Monmarché N. et al, 1999. On Improving Clustering in Numerical Databases with Artificial Ants. *ECAL'99 Proc. - European Conf. on Artificial Life,* pp. 626-635.

[22] Parpinelli R.S. et al, 2002. Data Mining with an Ant Colony Optimization Algorithm. *IEEE Trans. On Evolutionary Computing*, Vol. 6, no. 4, pp. 321-332.

[23] Ramos, V. and Abraham A., Swarms on Continuous Data, In *Proc. Congress on Evolutionary Computation,* IEEE Press, pp. 1370-1375, 2003.

[24] Ramos, V., Self-Organizing the Abstract: Canvas as a Swarm Habitat for Collective Memory, Perception and Cooperative Distributed Creativity, *1st Art & Science Symposium – Models to Know Reality*, pp. 59, 2003.

[25] Ramos V. et al, 2002. Self-Organized Data and Image Retrieval as a Consequence of Inter-Dynamic Synergistic Relationships in Artificial Ant Colonies. *Soft Computing Systems - Design, Management and Applications*, IOS Press, Vol. 87, Chile, pp. 500-509, 2002

[26] Ramos V. and Merelo J.J, Self-Organized Stigmergic Document Maps: Environment as a Mechanism for Context Learning, In *Proc.* of 1st Spanish Conf. on Evolutionary and Bio-Inspired Algorithms, pp. 284-293, 2002.

[27] Ramos, V., 2002, On the Implicit and on the Artificial – Morphogenesis and Emergent Aesthetics in Autonomous Collective Systems. *ARCHITOPIA – Art, Architecture, Science*, Portuguese Science and Technology Ministry, Lisbon, Portugal, chapter 2, pp. 25-57.

[28] Ramos V. and Almeida F., 2000. Artificial Ant Colonies in Digital Image Habitats - A Mass Behaviour Effect Study on Pattern Recognition". *From Ant Colonies to Artificial Ants - Procs. of the 2nd Int. Work. on Ant Algorithms*, Univ. Libre de Brussels, Belgium, pp. 113-116.

[29] Scott Camazine, Jean-Louis Denebourg, Nigel R. Franks, James Sneyd, Guy Theraulaz and Eric Bonabeau, 2001. *Self-Organiztion in Biological Systems*, Princeton Studies in Complexity, Princeton Univ. Press, Princeton and Oxford.

[30] Smith K.A. and Ng A., 2003. Web page Clustering using a Self-Organizing Map of User Navigation Patterns. *Decision Support Systems, Vol. 35, Issue 2*, pp. 245-256.

[31] Srivastava J. et al, 2000. Web Usage Mining: Discovery and Applications of Usage Patterns from Web Data. *SIGKDD Explorations*, Vol. 1, nº 2, pp. 12-23.

[32] Wang X. et al, 2002. Soft Computing Paradigms for Web Access Pattern Analysis. *Proc. of the 1st Int. Conf. on Fuzzy Systems and Knowledge Discovery*, pp. 631-635.

[33] Monash University Web site, 2002. *<http://monash.edu.au>*.